Building and Environment 72 (2014) 309–318

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

A data-driven method to describe the personalized dynamic thermal comfort in ordinary office environment: From model to application

Qianchuan Zhao [a], Yin Zhao [a,*], Fulin Wang [b], Jinlong Wang [a], Yi Jiang [c], Fan Zhang [c]

[a] Department of Automation and TNList, Tsinghua University, Beijing 100084, China
[b] Department of Building Science, Tsinghua University, Beijing 100084, China
[c] United Technologies Research Center (China) Ltd., Shanghai 200120, China

ARTICLE INFO

Article history:
Received 18 July 2013
Received in revised form
9 October 2013
Accepted 10 November 2013

Keywords:
Data-driven modeling
Personal thermal comfort
Least squares estimation
PMV
Personal energy saving potentials

ABSTRACT

Recent advances of information technology and the low cost of computing devices make it possible to collect end users' true thermal sensations at the building operation stage. This enables us to build personalized thermal comfort model from a new aspect. This paper proposes a data-driven method to describe the personalized thermal comfort in ordinary office environment. The model structure shows the condition of heat balance of human body, with four personalized coefficients estimated by on-line voting data. The adjustable coefficients provide the freedom to capture the personal differences in thermal comfort requirement. In contrast, the well-known PMV model is only an average model, which cannot reflect such differences. The model performance is evaluated by a field experiment study. A personal energy saving potential analysis is also presented as one of the applications. Both the experiment results and simulation results demonstrate the high accuracy of the data-driven model and the feasibility of the application in investigating personal energy saving potentials.

1. Introduction

With the rapid development of information technology, information acquisition system can be widely deployed in buildings. The low cost of sensing and computing devices makes it possible to measure almost everything we want in buildings and carry out learning and calculating on line. This will bring us a large amount of opportunities to make the building environment more comfort and efficient. This paper focuses on describing the personalized thermal comfort using information technology.

Although currently there are many researches in modeling human comfort, most of them actually focus on the average thermal comfort model. The well-known work, Predicted Mean Vote-Predicted Percent Dissatisfied (PMV-PPD) model [1,2] is based on a large data set. It establishes an average mapping from the environmental factors (air temperature, radiant temperature, relative humidity, air velocity) and personal factors (clothing level, metabolic rate and activity level) to the 7-level comfort value scale. The following validation, modification and extension of the PMV-PPD model by Humphreys MA [3] use the ASHRAE (American Society of Heating, Refrigerating and Air Conditioning Engineers) database for large groups of persons. They do not deal with the personal comfort. Besides the researches mentioned above, other related researches have similar limitations. Examples might be the models which are based on the human body physiology, such as Gagge AP's two-node (core and skin) model [4,5]. This model represents the body as two concentric cylinders, with the thermoregulatory and blood flow considered. However, it is based on the general physiology of human body without personal factors. The commonly used ET* and SET* [6] are based on it. Similar work includes the Stolwijk JAJ [7] multi-segment mathematical model of human body, which divides the human body into finer parts, and Zhang et al. [8]'s sensation and comfort model for human segments and the whole-body. de Dear RJ and Brager GS [9], Humphreys MA [10], etc. proposed adaptive models of comfort, which give a relationship between the indoor comfort temperature and the outdoor mean temperature and reveal the adaptivity of human comfort. They are still based on a large amount of field studies and statistics. Many of those works have been written in the ASHRAE standard and play important roles in building construction design, HVAC system capacity design and operation management.

However, in practice, the individual differences in thermal preference often incur dissatisfactions, especially in the operational stage. This indicates that the average model, such as PMV, may have

* Corresponding author. Tel.: +86 10 62792491; fax: +86 10 62796115.
E-mail address: zhaoyin09@mails.tsinghua.edu.cn (Y. Zhao).

0360-1323/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.buildenv.2013.11.008

bias in predicting the thermal comfort for a small number of people. The PMV-PPD itself also points out this by stating that even the environment satisfies PMV = 0, there is still 5% users who do not feel acceptable. Even if in Comfort category C of ISO 7730 [11], where comfort requirement extends to [−0.7, 0.7], there is almost 16% dissatisfied percentage. When PMV = 1 the dissatisfied percentage is up to 30%. Although individual differences in thermal comfort have been noted, such as differences in young women [12] and between genders [13], little modeling work has been done on it. The challenge of building personalized comfort model is how to determine the model structure in a clear way so that it can capture the thermal preferences for individuals.

The aim of this work is to provide a rationale and structural model as well as the learning method to describe the personalized comfort in ordinary office environment. In this paper, we develop a data-driven learning method for the personalized thermal comfort using the feedback data from information acquisition system in real office environment. Different from the traditional chamber research, field study and statistics analysis for large groups, we derive a rationale structure based on body heat balance, leaving the personalized coefficients to be estimated using the measured personalized voting data. Even when the thermal sensation changes with time, the update scheme could assure the tractability of the model. It avoids the detailed specification of the personal factors, such as clothing level, activity level, etc., which are usually needed in existing comfort model but hard to determine in practical application. The current work is restricted to the personalized thermal preference and could be a first step for many following researches related to the personalized comfort, such as the personalized energy consumption and comfort tradeoff, the personalized comfort control algorithm and even the multi-person office comfort control, etc. We only present the analysis method of individualized trade-off between energy consumption and comfort as an illustrative example of the applications. The method is based on the proposed personalized comfort model with interchangeable energy models.

The rest of the paper is organized as follows. After the introduction, the proposed learning model is introduced in Section 2. Then we present an application of the model – personalized energy saving potentials analysis in Section 3. Experimental results and case study result of personal energy-saving analysis are in Section 4 to demonstrate the methods. We conclude our paper in Section 5.

2. Personalized dynamic thermal comfort

We first derive the parameterized structure of the personalized thermal comfort model. Then the coefficients estimation method is given. The model is a data-driven gray-box model, whose structure is based on human body heat balance equation, with the personalized coefficients to be identified using the personalized feedback voting data.

2.1. Heat balance of human body

The key factors of the heat balance of a human body are the rate of metabolic heat production M (W), the rate of mechanical work W (W), convective sensible heat loss from skin C (W/m^2), radiant sensible heat loss from skin R (W/m^2) and the rate of evaporative heat loss from skin and respiration E (W/m^2). Fanger [1] proposed that if a human body is in a thermal steady state, the human body keeps a thermal balance state by the thermoregulatory system as shown in Equation (1),

$$M - W - R - C - E = 0 \tag{1}$$

where the heat loss equates the heat production. In the PMV framework, assuming the skin temperature is around its neutral level, the thermal L load of human body is defined as the difference between heat production and heat loss as shown in Equation (2).

$$\begin{aligned} L &= (M-W) - (R+C+E) \\ &= (M-W) - \Big(3.96\times10^{-8} f_{cl}\Big[(t_{cl}+273)^4 - \big(\overline{t_r}+273\big)^4\Big] \\ &\quad + f_{cl}h_c(t_{cl}-t_a) + 3.05[5.73 - 0.007(M-W) - P_a] \\ &\quad + 0.42[(M-W) - 58.15] + 0.0173M(5.87 - P_a) \\ &\quad + 0.0014M(34 - t_a)\Big) \end{aligned} \tag{2}$$

where f_{cl} is the clothing area factor, i.e. the ratio of clothing surface area to the surface area of nude human body, t_{cl} is the clothing surface temperature (°C). Note that the calculation of t_{cl} needs to solve the equations $C = f_{cl}h_c(t_{cl} - t_a)$, $t_{cl} = t_{sk} - I_{cl}(R + C)$ and $R = 3.96\times10^{-8} f_{cl}\,[(t_{cl}+273)^4 - (\overline{t_r}+273)^4\,]$. If we substitute R and C into $t_{cl} = t_{sk} - I_{cl}(R+C)$, than we get a nonlinear equation of t_{cl} given t_{sk} and I_{cl}. Thus it can be solved using numerical method. More details can be found in Ref. [6] $\overline{t_r}$ represents the average radiant temperature (°C), t_a is the room air temperature (°C), h_c is the convective heat transfer coefficient (W/(m^2 K)) and P_a is the water vapor pressure (kPa). The mean response of a large group of people corresponding to the ASHRAE thermal sensation scale of [−3, 3], i.e. the Predicted Mean Vote (PMV), can be described using the thermal load L using the following Equation (3).

$$PMV = [0.303\exp(-0.036M) + 0.028]L \tag{3}$$

2.2. Model structure

Everyone has different metrics to the environment in their minds. Those differences might result from many factors, such as their living areas, ages, genders and living habits. In addition, the psychological factors indeed have influences on human expressions of their thermal perceptions. However, the most important factor that determines the sensation is the thermal balance of human body. Thus the personal model structure is based on it. The inputs of PMV can be categorized into two types: the environmental factors, i.e. air temperature, mean radiant temperature, vapor pressure, etc., and the personal factors such as the clothing level, metabolic rate, mechanical work, etc. To describe the personalized thermal comfort, we keep the environmental factors $(t_a, \bar{t}_r, h, v)$ unchanged and aggregate the personal factors together as the personalized coefficients to be identified. Then a parameterized equation in the form of the environmental factors can be obtained after some mathematical manipulations. We name it personalized thermal vote (PTV) and write it as

$$PTV = m_0 + m_1P_a + m_2t_a - m_3(R+C) \tag{4}$$

where m_i, $i = 0, 1, 2, 3$ are personalized coefficients to be identified, P_a is dependent on t_a and h, the relative humidity and $R + C$ depends on $t_a, \bar{t}_r, v$.

Several issues should be noted. First, Equation (4) is essentially a direct derivation of PMV Equation (3). That means if we substitute $m_0 = (0.303e^{-0.036M} + 0.028)(0.4523M - 0.6014W + 6.95)$, $m_1 = (3.05 + 0.0173M)(0.303e^{-0.036M} + 0.028)$, $m_2 = 0.0014M(0.303e^{-0.036M} + 0.028)$ and $m_3 = (0.303e^{-0.036M} + 0.028)$ into Equation (4), Equation (4) is the same with Equation (3), i.e. the PMV model. Second, some simplifications are carried out in term $m_3(R + C)$. As we know, $R + C$ is decided by the air temperature, radiant temperature, clothing level and the air velocity. They are highly coupled through nonlinear equations and unable to be decoupled into personal factors and environmental factors. Here, we calculate $R+C$ with the clothing level as constant (1.0 clo) as a base line. Then $R+C$ only depends on the environmental factors, i.e. the air temperature, mean radiant temperature and air velocity. The coefficient m_3 is a modification term considering the possible variations of $R+C$ caused by the personal factors, such as the clothing level. We further assumes $v \leq 0.1$ m/s because the air velocity in office environment is generally in a low and constant level. Then $R+C$ is dependent on $t_a, \bar{t}_r$. Third, the coefficients to be identified in our model are based on the personal voting data. This is rational because those coefficients only depend on the personal factors in the framework of the heat balance of human body. When human thermal sensation changes, user will have different thermal sensation votes. Receiving the new data, the coefficients will be updated to capture these changes. We do not exploit the reasons of the sensation changes because they are related to many factors uncontrollable. The proposed model actually concludes all those possible reasons to the coefficients to be estimated. Fourth, human sensation might also have some relationships with the location relative to the supply diffuser or windows. However, we consider a relative steady environment, i.e. subject's position is relatively fixed. The differences between different subjects might be caused partly by the position, but for one person, if the influence is expressed through the user's vote, the parameters will be update to capture it. Besides those above, the application of the model is restricted to the steady state comfort because the basic principle behind the derivations is the steady heat balance of human body.

By introducing the time index in Equation (4), we finally obtain the Personalized Dynamic Thermal Comfort (PDTC) model

$$PDTC(k) = m_0(k) + m_1(k)P_{a,k} + m_2(k)t_{a,k} - m_3(k)(R+C)_k \tag{5}$$

where $m_i(k), i = 0, 1, 2, 3$ are on-line tuned at each time step k when the occupant votes his/her thermal sensation. The time step could be one hour or an even longer interval because the model is only applicable for the change of thermal sensation in steady state. Transient thermal sensation might involve more detailed physiological analysis and is beyond the scope of this paper.

2.3. Coefficient estimation

Given the structure in Equation (4), (5) and the personalized thermal sensation voting data, the coefficients of the PDTC can be estimated. For different usages, different estimation schemes could be utilized. One of the estimation schemes is to estimate the coefficients using the data of one certain period to analyze the personalized thermal comfort preference in it. In this scheme, each period should have relatively constant personal factors so that the difference of the coefficients between different periods could reflect the thermal perception variations. The coefficients can be estimated using the commonly used least squares method as below:

$$\min_{m_0,m_1,m_2,m_3} \sum_{i=1}^{N} (\tilde{y}(i) - PDTC(i))^2 \tag{6}$$

where $\tilde{y}(i)$ is the actual vote and it is in the range $(-3, 3)$ same with PMV, and N is the total number of votes for one subject in the period.

Another estimation scheme is the weighted least squares estimation considering the dynamic property of the human thermal perception. It is especially applicable for the online learning and control. The coefficients at time k for a certain subject are calculated using the following optimization:

$$\min_{m_0(k),m_1(k),m_2(k),m_3(k)} \sum_{i=1}^{k} \gamma(i)(\tilde{y}(i) - PDTC(i))^2 \tag{7}$$

where $\gamma(i) \in (0, 1)$ is the weighting factor at time i. Generally, the weighting factor is chosen as λ^{k-i+1} for dynamic system, where λ is the forgetting factor, i.e. putting more weights on the recent data and forgetting the historical data gradually. Optimal selection of the forgetting factor has been derived in Ref. [14]. It needs the statistics of measurement noise and process noise, which are unknown in our case. Because the model is to capture the possible changes of thermal sensation in a relative long term, the forgetting factor is recommended to be close to 1. Combining the interval of the votes and the forgetting factor, one could know how the historical data is weighted. For online implementation, the solution of Equation (7) can be implemented in a recursive way, i.e. Recursive Least Squares algorithm [14]. It sequentially estimates the coefficients as the data sequentially available. The initial coefficients can be estimated using the least squares estimation in an initial period (such as one week's data). The flowchart of the algorithm and the related equations are shown in Fig. 1 and more details can be found in Ref. [14].

3. Personal trade-off between energy and comfort by the model

Having developed the data-driven personalized comfort model, we present one of the applications, using the proposed model to analyze the personalized trade-off between energy and comfort. The energy consumption of buildings takes up to 40% of the whole energy consumption [15]. Many attentions have been paid on how to achieve the energy saving while maintaining comfort. However, due to the lack of personalized comfort model, the analysis of energy consumption based on the comfort model is only for the average model, generally PMV. Typical work involves the model predictive control using the linear combination of comfort and energy cost as objective function [16] and the multi-objective optimization using thermal comfort and air quality as the comfort index and heat/cooling load of the HVAC as the energy index [17]. Zhong K et al. [18] analyzed the energy saving potentials in different buildings in different climate areas in China. The comfort model was the adaptive comfort model [9]. In this section, we present a framework to analyze the personal trade-off between energy and comfort using PDTC. The analysis is essentially a further review of the PDTC model in personal difference, i.e. the personalized comfort margin to compromise. The general idea for this analysis is minimizing the energy consumption under the condition that the personal thermal sensation should be acceptable. It is an extension of the method in Ref. [18] with the proposed personal comfort model incorporated and the energy model interchangeable.

3.1. Energy consumption estimation

To analyze the personalized energy consumption and comfort, an energy consumption model will be needed. Similar to Ref. [18], the key bridge between the proposed personalized comfort model and the energy consumption is the set-points of the thermal environment of the room, i.e. air temperature, relative humidity, etc. The heating/cooling load under certain environment could be used to study the energy performance. Given a certain room with its location, climate information, and the occupancy information, the heating/cooling load would be a function of air temperature,

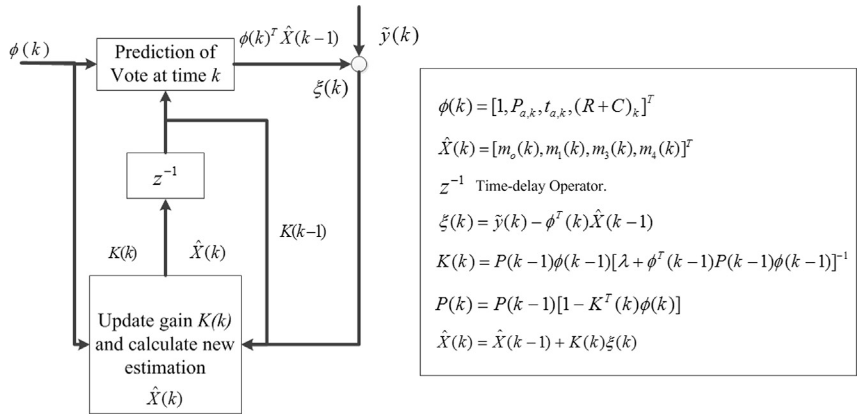

Fig. 1. Flow chart of the Recursive Least Squares Estimation and the related equations.

relative humidity. Without loss of generality, we denote the heating/cooling load as $Q_{all}(t_a, h_a)$. $Q_{all}(t_a, h_a)$ could be of any form, such as the steady transfer estimation, software simulation estimation (such as DeST [19], EnergyPlus [20] or DOE2 [21]), or a more advanced energy model for specific HVAC systems. Note that here we only use the thermal load as the representation of the energy consumption. For the detailed energy consumption, one should consider the detailed types of devices and equipment. Different devices with different efficiencies will have variable energy consumptions under the same heating/cooling load. We do not consider the device-dependent energy consumption in this paper.

3.2. Set-point optimization

The optimal set-point is obtained using a constraint optimization formulation. The objective function is the heating/cooling load $Q_{all}(t_a, h_a)$ mentioned above, which is the function of the temperature and relative humidity of the room. One of the constraints is the human comfort. For the personalized comfort, it should keep the PDTC in the range of $(-threshold, threshold)$. That is the absolute value of PDTC is inside an interval around 0, e.g. $(-0.5, 0.5)$. This assumes the symmetry of PDTC in cold and hot, which is similar to PMV. Also, some more restricted constraints in air temperature and relative humidity (t_{down}, t_{up}), (h_{down}, h_{up}) should be given to assure they are in a reasonable range (h_{down}, h_{up} are the lower bound and upper bound of the relative humidity, and t_{down}, t_{up} are those of temperature). Thus, the optimal temperature and relative humidity at time k of the room are determined as

$$\begin{gathered} Min_{\{t_{a,k},h_{a,k}\}} Q_{all}(t_{a,k}, h_{a,k}) \\ s.t.\left|m_0(k) + m_1(k)P_{a,k,} + m_2(k)t_{a,k} - m_3(k)(R+C)_k\right| \le threshod \\ h_{a,k} \in (h_{down}, h_{up}), t_{a,k} \in (t_{down}, t_{up}) \end{gathered} \tag{8}$$

In our estimation, $Q_{all}(t_{a,k}, h_{a,k})$ considers both heating and cooling load together. For winter case, the heating load is positive and the minimization means choosing the set-points that lead to the minimal heating requirement in the room. For summer case, the cooling load is positive and the minimization means choosing the set-points that lead to minimal cooling requirement in the room (in our following case study in Section 4, we take winter case as an example). Considering that the optimal temperature set-point variations are very small, the dynamic influence to heating/cooling load of temperature set-point change is ignored. While the thermal dynamics of room air heat gain are considered by using cooling load factor method to calculate heating/cooling load. If considering more detailed historical dependence and dynamics, the optimal temperature and relative humidity at time k are highly dependent on the dynamics of the thermal load. The optimization formulation for dynamic system might be needed. Different orders of dynamics will need different optimization techniques and simplification methods and those are beyond the scope of this paper.

If $Q_{all}(t_{a,k}, h_{a,k})$ has an analytical formula, the problem falls into the field of nonlinear optimization, where the local optima can be obtained using mature optimization methods, such as the gradient-based method or the interior-point method. Note that the solution might be local-optimal, thus several random initial solutions need to be tried to avoid getting trapped in a local minimum. For a simulation model, where $Q_{all}(t_{a,k}, h_{a,k})$ is evaluated using simulation, simulation based optimization techniques can be used to solve it. Readers could referred to Ref. [22] for more information.

3.3. Personalized energy saving metrics

We present two energy saving potential metrics to analyze the personalized tradeoff between the energy saving and comfort. These metrics are based on the results of optimization stated above. A baseline is first introduced, which is the optimization result of Equation (8) with PMV model as comfort constraint shown in Equation (9).

$$\begin{gathered} Min_{\{t_{a,k},h_{a,k}\}} Q_{all}(t_{a,k}, h_{a,k}) \\ s.t.|PMV| \le threshold \\ h_{a,k} \in (h_{down}, h_{up}), t_{a,k} \in (t_{down}, t_{up}) \end{gathered} \tag{9}$$

It is the average result and then the personalized results could be compared with it. The first metric is the personal energy savings relative to the PMV results, i.e.

$$RPE = \left(Q^*_{PDTC}(t_a, h_a) - Q^*_{PMV}(t_a, h_a)\right) \Big/ Q^*_{PMV}(t_a, h_a) \times 100\% \tag{10}$$

where $Q^*_{PDTC}(t_a, h_a)$ is the optimal objective function of problem (8) when the comfort constraint is PDTC, and $Q^*_{PMV}(t_a, h_a)$ is that of PMV for problem (9). Here we omit the time index k. Both$Q^*_{PDTC}(t_a, h_a)$ and $Q^*_{PMV}(t_a, h_a)$ can be the results at certain time k or in certain period. This metric is an intuitive result presenting whether and how much the user consumes the energy relative to the average level, which is similar to the metric defined in Ref. [18].

The second metric is the personalized sensitivity analysis of the energy consumption versus thermal comfort. The personalized energy saving sensitivity is defined as follows

$$S_{PDTC,\Delta} = \left| Q^{*}_{PDTC,threshold} - Q^{*}_{PDTC,threshold+\Delta} \right| / Q^{*}_{PDTC,threshold} \times 100\% \quad (11)$$

It is the relative heating/cooling load decrease when the comfort constraint relaxes from *threshold* to *threshold* + Δ. The sensitivity indicates how much the energy saving can be achieved when the user compromises in the comfort requirement.

The above two metrics could be used to characterize the personal energy-saving potentials. The relative decrease of energy consumption gives the baseline of the energy saving for the person and the sensitivity gives the margin of further energy savings by a little compromise in thermal comfort. When one obtains the personalized coefficients of PDTC model, the energy saving potentials for the user can be analyzed using the above method with a proper energy model.

4. Experiment and results

In this section, we present the experiment results of the above methods and the tradeoff analysis results in a case study. These results demonstrate the effectiveness and applicability of the proposed model and its application in personalized energy saving analysis.

4.1. Coefficient estimation using on-line survey data

4.1.1. Experiment setup

To demonstrate the feasibility and performance of the proposed data-driven model, a simple experiment – an online voting survey, was carried out. The experiment lasted from November 2009 to January 2010, i.e. the fall and winter season in Beijing, with the monthly average outdoor temperature changing from 11 °C to 0 °C. The experiment office is 5 × 12 m^2, with 5 graduate students and 4 professors and researchers working in it as subjects. They were dispersed so that they did not influence each other's perception (Fig. 2). There were 5 sets of sensors (air temperature sensor, mean radiant sensor, air speed sensor, humidity sensor) distributed in the office to assure there was at least one set of sensors nearby for each subject to measure the environment around them. The HVAC system in the office is Fan Coil Unit, located at two diagonal corners. They were controlled by the users as normal and the environment of the office varied in a reasonable range. The advantage of this setting is that the experiment is nearly the same to the real office condition and subjects' behavior will not be influenced by the "experiment". The histogram of the mean air temperature of the office during the experiment is illustrated in Fig. 3.The air speeds around subjects were less than 0.1 m/s. The room was mechanical

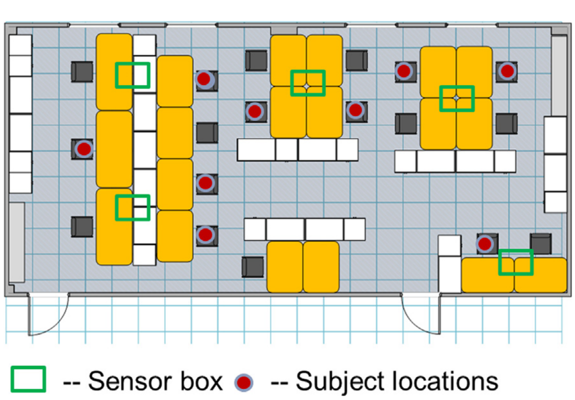

Fig. 2. The deployment of the sensor boxes and subjects in the office.

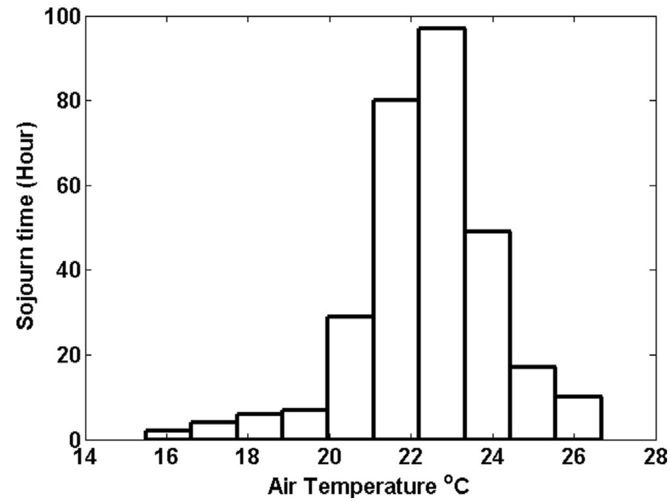

Fig. 3. Histogram of the mean air temperature during the experiment.

ventilating (at least 30 m^3 for each person per hour) to guarantee the air quality.

Occupants in the room were asked to keep sedentary and to do normal office work. Additionally, they voted in their computers through the dedicated software that popped up every one hour during their office time. The thermal sensation vote panel was designed as a scrollbar, with five sensation marks, which were hot, warm, neutral, cool and cold. The neutral perception is associated with 0, warm with 1.5 and hot with 3 and users can vote continuously between −3 and 3. Users can vote continuously from −3 to 3. We only use the five marks to guide the user's vote because it is easier to understand than a set of more complicated labels. Choosing the range (−3, 3) is also the same as PMV scalar, which makes it comparable to PMV results. The subjects had been trained to be familiar with the voting interface before experiment. In order to avoid the less concentration on the vote, we only required the subjects to vote on thermal sensation every one hour. There are 2679 votes collected, averagely around 300 votes for each subject and only 7–8 times in one day. So the voting is expected not to interrupt the subject too much. Fig. 4 gives an overview of each subject's vote values, including the mean value and standard deviation. Generally, the mean values of each subject's votes are different. Since they experienced similar environment, the difference of the overall statistic indicates their difference in thermal

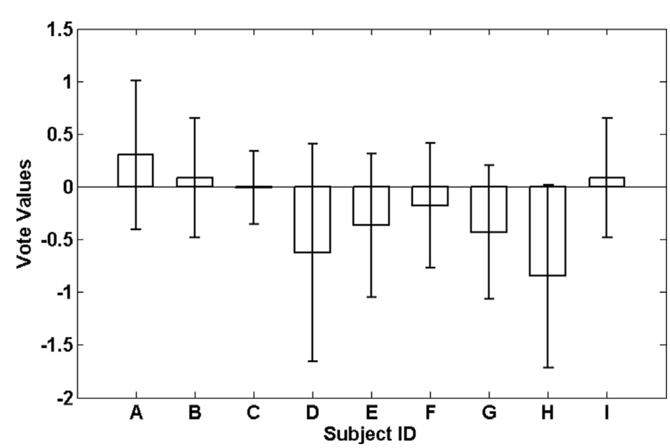

Fig. 4. Overview of each subject' vote during the experiment, the bars show the mean value and the intervals show the standard deviation of the votes.

preference. Fig. 5 shows a typical day's vote data and environment measurement. The first vote of each day for each person is ignored during the calculations since at the beginning of the day, the occupant' first vote generally bears more noisy information because of the unstable physiological levels caused by outside movement on their way to office, such as walking or riding. We are not trying to describe the transient thermal comfort.

4.1.2. Results

We first estimate the coefficients using the least squares estimation for the data in the each month for each subject. Table 1 gives the estimation results for four of the subjects. For each subject, we used the data in each month to estimate the coefficients. The input of the estimation is the personal actual vote and environment factors, i.e. the air temperature, mean radiant temperature and relative humidity measured by sensors. Those environment factors were used to calculate P_a, t_a, $R+C$ and then the parameters were estimated using Equation (6). Differences can be seen between different subjects and months. A reference value for PMV is also presented. The PMV corresponding value is calculated using the definition in Section 2.2. Fig. 6 gives a relative small time scale coefficient estimation result using the recursive estimation method. The estimation updates the results each time a new sensation vote is received. As stated in Ref. [14], the RLS algorithm memory length is measured using $1/(1-\lambda)$. The forgetting factor value should be determined case by case. In our experiment setting, the voting interval is 1 h and each subject votes averagely 8 times in one day. We set the forgetting factor as 0.98, which means the algorithm uses around $1/(1-0.98) = 50$ voting data (memory length) to estimate factors, which is around one week's historical records. Smaller selection of forgetting factor will reduce the memory length of algorithm and increase the sensitivity of the estimated parameters to noise. Thus it is also a tradeoff between stability and sensitivity. We plotted each parameter in one figure so the differences between subjects can be seen clearly. We still take the four subjects as an example and the four parameters are separately drawn in the four subfigures to explicitly show the differences. As expected, we could see a detailed difference between the four individuals and detailed changing of the coefficients over time. Note that some dramatic changes of the coefficients as we have labeled out in Fig. 6 generally correspond to the outdoor weather dramatic changes. For example, around 17th, December, 2009, the outside lowest temperature dropped down significantly from -5 °C to -14 °C, at the same time, the averaged wind speed jumped from 7 km/h to 26 km/h. A relative sharp change in coefficients could be found in almost all the subjects. Similar cases happened in 5th and 12th in January, 2010. The relationship between the thermal sensation and outdoor weather has been studied by de Dear and Brager [9] and Nicol et al. [23]. The dynamic estimation could reflect this relationship. To give a more illustrative comparison between our model's predicted comfort temperature and the adaptive comfort model by de Dear et, al., we choose the average outside temperature in the three months, i.e. (11 °C, 4 °C, 0 °C respectively) and use the estimated parameters shown in Table 1 to calculate the comfortable temperature for each subject. We set the relative humidity as 40% and further assume the air temperature and mean radiant temperature are equal to reduce the dimension of search. (Actually in our experiment, the difference between the two temperatures is less than 0.8 °C). The results are plotted over the adaptive model results in Fig. 7. We can see that most of subjects show a similar trend of adaptive comfort preference with the outdoor temperature while different subjects appears in different position reflecting the individual difference. This indicates that our proposed model is able to capture the adaptive comfort phenomenon in personal level.

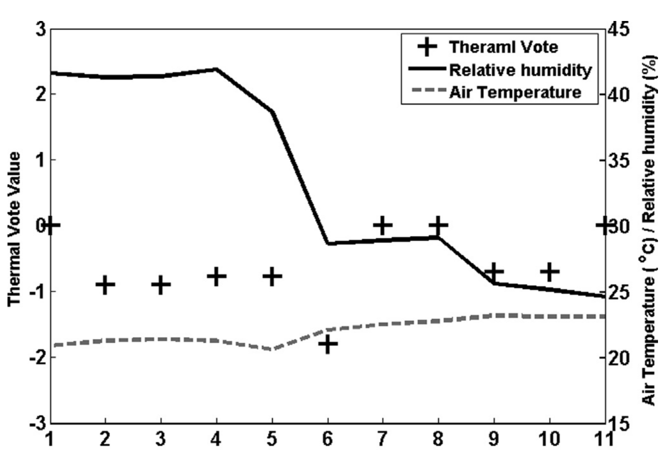

Fig. 5. A typical day's voting data, the x-axis is the vote sequence in one day and the interval time is about one hour.

Table 1
Parameter estimation results month by month.

Subject	Months	m_0	m_1	m_2	m_3
A	Nov.	5.917	1.279	−0.242	0.211
	Dec.	3.185	0.307	−0.17	0.007
	Jan.	14.789	2.996	−1.777	2.227
B	Nov.	1.735	0.272	0.0515	0.425
	Dec.	2.088	0.051	0.0119	0.311
	Jan.	2.336	−0.14	0.0036	0.296
C	Nov.	6.67	0.687	−0.168	0.529
	Dec.	5.93	1.478	−0.177	0.479
	Jan.	5.385	1.984	−0.164	0.48
D	Nov.	−1.358	−0.466	0.07	0.023
	Dec.	−4.423	1.112	0.135	−0.053
	Jan.	9.663	0.665	−0.242	0.692
PMV	$M = 58$ W/m^2, $W = 0$, Icl = 1 clo	2.18	0.264	0.0053	0.0651

Table 2 shows the quantitative evaluation of the accuracy for the recursive estimation. The quantitative metrics are the Bias and the Mean Square Errors (MSE). To do it, every third of the votes are chosen as the validation data (33% of all the data) and the rest are training data (67% of all the data). For the training data, the coefficients are updated and the regression errors are obtained. For the validation data, we use the current learned model to predict the human sensation vote and get the difference between true vote and predicted vote, i.e. the prediction error. The first four columns of Table 2 are the Bias and MSE results for the two types of errors. They reflect the performance of the model in regression accuracy and predictability. The prediction errors using PMV in personal comfort are also presented as a base line, on one hand, to show the limitation to use PMV to describe personal comfort, on the other hand, to demonstrate the advantage of the proposed PDTC model. The clothing level when calculating PMV was 1.0 clo and other inputs were the measured values at each vote. Overall, both the regression error and prediction error of PDTC for each user are smaller than PMV, although the prediction errors are slightly higher than the regression errors. It indicates that the proposed method has obtained personalized coefficients for each user with a higher accuracy. (Even if we merge the 9 subjects' voting data together to form the mean vote, the average bias of using PMV model is still around 0.5, indicating the limitation of PMV's prediction for personal and small group of persons). Specifically, for the subject A and E, the MSEs of PDTC are much smaller compared to the base line PMV. For subject B and I, the bias of PDTC are also

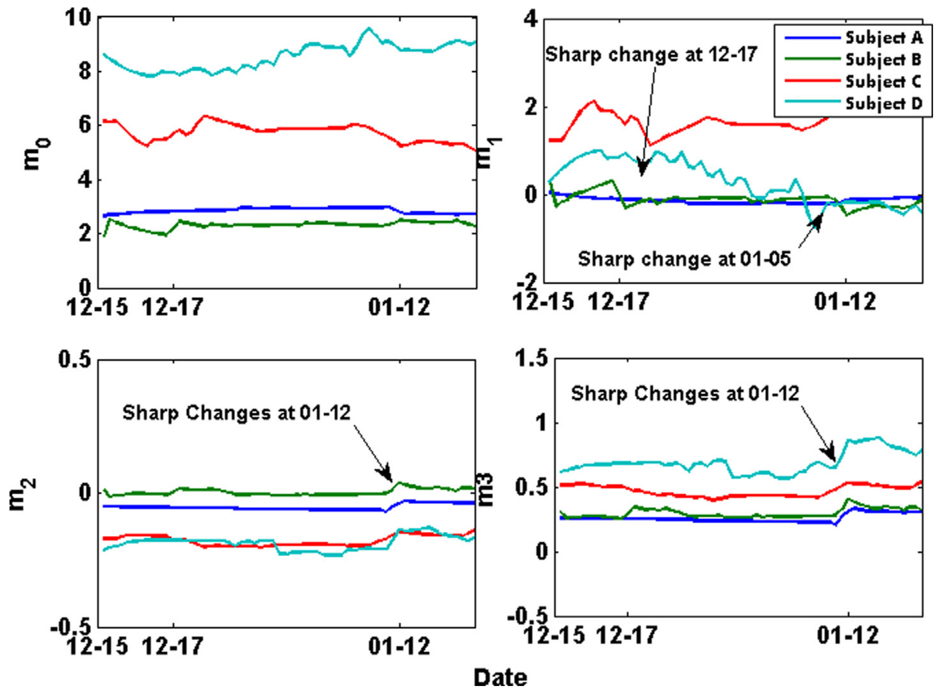

Fig. 6. Recursive regression results of four typical subjects during their experiment periods. Each subfigure shows one of the estimated parameter results for the typical four subjects.

much smaller. Of course, although some MSEs of PDTC are a little greater than those of PMV, such as H, the bias is much smaller.

Fig. 8 further gives the results of the cross-correlation coefficients between the prediction error sequence and the historical input sequences. It is to test the tractability of the model and algorithm [24]. In Fig. 8, R_{ξ,ϕ_i} represents the cross-correlation coefficients of prediction error ξ and inputs sequences $\phi_1 = 1$, $\phi_2 = P_a$, $\phi_3 = t_a$, $\phi_4 = R + C$, which are already defined in Fig. 1. The independence between the prediction error sequence and historical input sequence means that the information in the prediction error is independent of the model and parameters, reflecting that the model has a satisfying performance considering the tractability of the dynamics [24]. The absolute values of the coefficients in the results are less than 0.3 at different time offset, which is a relatively low level, indicating that the historical inputs have a weak correlation with the residuals.

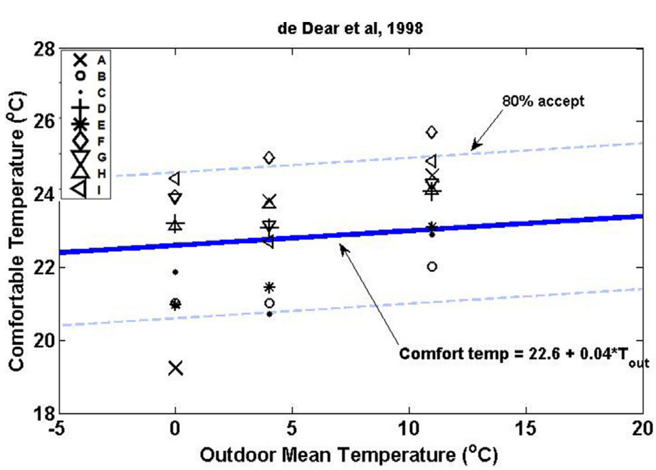

Fig. 7. Subjects' comfort temperature variations and comparisons with adaptive comfort model.

4.2. Results of energy saving potentials – a case study

4.2.1. Study case setup

In this case study, the heating load for a single office room in a building in Beijing during winter season was used as the energy estimation. The room is about 6×6 m^2 with 3-meter high. The east wall is external and the others are internal. The windows are east orientated. Fig. 9 shows the key factors considered in the estimation of heating load based on mass and energy balance [25]. They are based on the following major assumptions and simplifications: 1. The air in the room is well mixed. 2. Solar radiation is not considered because the solar heat gains do not change with different room air temperatures [25]. 3. We use hour-by-hour heating load estimation to approximate the load profile at different hours. Considering that the optimal temperature set-point variations are very small, the transient behavior caused by temperature set-point change between the time steps is not considered [25]. However, we still use cooling load factors described in Ref. [26] to calculate the heat gains of the heat source in the room (see Table 3) 4. Constant outside air volume that meets the healthy requirements is supplied during the daytime (8:00–21:00). We also assume that the occupant, whose comfort requirement is modeled by PDTC model with coefficients obtained from the experiment, is in the office. Although it might need further experimental validation whether the coefficients are still applicable when the user is actually in the room, we still use it because it is only a simulation investigation of the proposed model in analyzing the energy saving potential. Table 3 shows the basic parameters of the room to estimate the heating load and Fig. 10 illustrates the outdoor weather condition [27], heat gain of the room due to the occupants, lights and appliance in office room in the case study. The occupant is assumed to be in the room from 8 AM to 11 AM and 1 PM to 5 PM. The room heat gains are calculated using cooling load factor method and thus at 12 AM and after 6 PM, there are still heat gains although there is no occupant in the room. Detailed heating load equations could also be found in Ref. [27].

Table 2
Mean square error and bias of PDTC and PMV, R-: regression; P-: prediction.

Subject	PDTC R-MSE	PDTC R-Bias	PDTC P-MSE	PDTC P-Bias	PMV P-MSE	PMV P-Bias
A	0.7230	0.009	0.9703	0.07	2.4916	0.06
B	0.5319	−0.015	0.5980	−0.034	1.2999	0.575
C	0.1442	−0.058	0.1363	0.026	0.5885	0.058
D	0.5182	0.064	0.5356	−0.05	0.4327	0.272
E	0.7860	0.064	0.9019	0.25	3.4994	−0.14
F	0.2860	0.036	0.2684	0.0214	0.713	−0.047
G	0.3607	−0.061	0.3634	0.1370	0.4633	−0.26
H	0.7167	−0.087	0.8088	−0.139	0.6777	0.249
I	0.2371	−0.025	0.2209	0.023	0.264	0.932

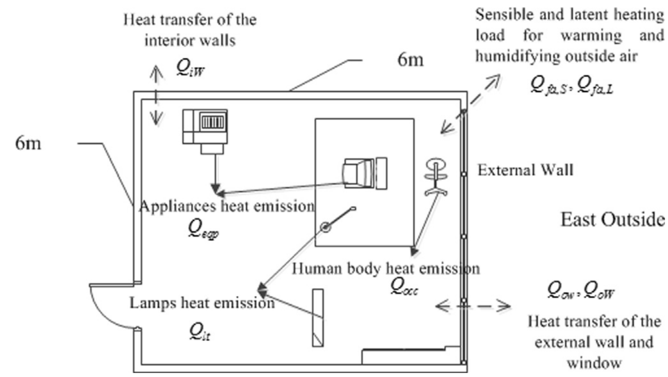

Fig. 9. The heating load estimation illustration of the single office room.

4.2.2. Results

Fig. 11 shows the heating load results of 4 subjects by solving the optimization (8) using the experiment results of PDTC as the comfort constraints in one day. Matlab optimization solver with the interior-point method was used. In order to cover all the possible comfort ranges for different individuals, the temperature range in formulation (8) is (18, 27), which is larger than the result in ASHRAE standard [28]. Optimized set-points using PMV are also illustrated in Fig. 11, with clothing level as1 clo and metabolic rate as 58.2 W/m^2 (1 met). The upper parts of Fig. 11 show the optimized set-points of indoor air temperature and relative humidity. For subject A and B, the optimal temperature is about 3–4 °C lower than that of PMV model, which means that they preferred cooler environment than the average. It should be pointed out that the level of clothing has impact on the model coefficients and further on the analysis results. We consider this factor indirectly by measured voting data. For example, in our experiment, subject A did experience temperature 18 °C and voted comfort in our experiment records. This is a result of heavy clothing .Thus the comfort temperature is lower. For subject C and D, the difference is relatively small, less than 1 °C, meaning that they are more close to the average. The difference in relative humidity is not significant, generally about 10%. Also, the optimized set-points sometimes vary at different time in one day due to the adaptive capability of PDTC to the individual's perception variations at different time. The hourly heating loads under the optimized set-points are shown in lower parts of Fig. 11. It is obvious that different user's thermal requirements lead to different heating loads. In this case study, subject C and D need relatively more heating compared with subjects A and B, which is in accordance with the results that subjects A and B prefer much cooler environment in winter.

Fig. 12 is the relative energy saving and sensitivity results for all the 9 subjects in this study case. The results are averaged for all the available voting data during the online survey process for each user. The two metrics are plotted. Most of the subjects in this experiment have relatively lower energy consumptions compared with the PMV-decided set-points. The heating load requirement for subject H is relatively higher than others. Also, the sensitivities for the subjects vary averagely from 1% to 6%. Although there are only 9 subjects in this case study, which is a small sample size, we could still obtained some observations on the personal difference in

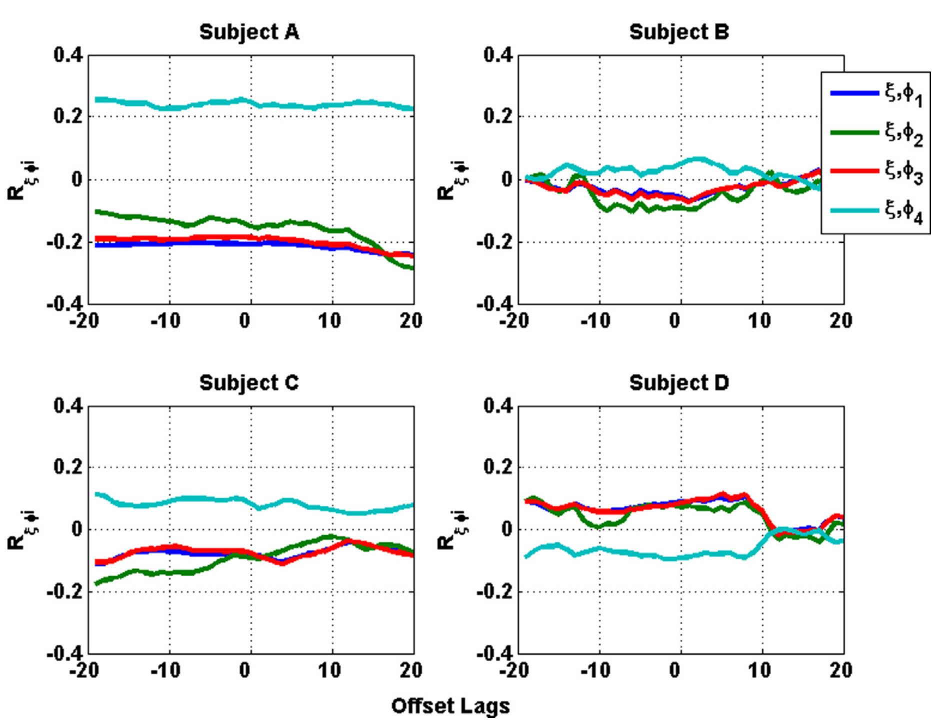

Fig. 8. Cross-correlation coefficient between residuals and past inputs for four of the subjects.

Table 3
Main parameters of the room in the heating load estimation.

Occupancy time	Number of occupant	Fresh air volume	Window size	Height of the room
8:00–21:00	1	30 m^3/h	18 m^2	3 m
Thermal transmittance of walls	Thermal transmittance of window	Heat emission of lighting	Heat emission of appliances	Temperature difference between adjacent rooms
0.8 W/m^2 °C	2.3 W/m^2 °C	450 W	800 W	3 °C

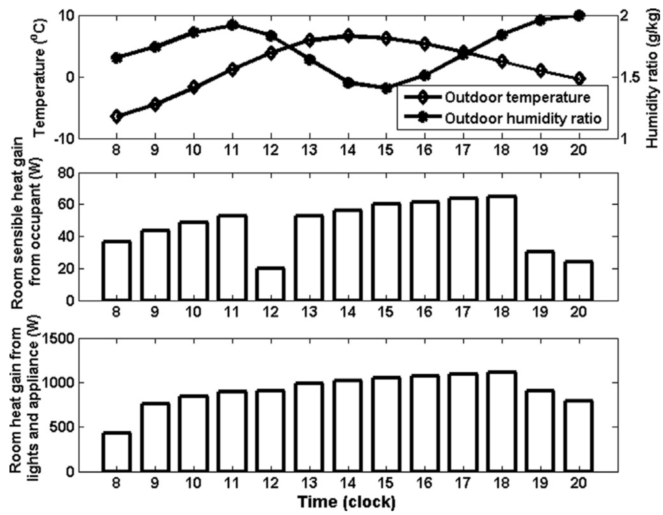

Fig. 10. Outdoor air conditions and room heat gain due to occupants, lights and appliances in simulation.

energy saving tradeoff. The sensitivities for high-energy users are generally higher than low-energy users. The subjects who have both higher energy consumption and higher sensitivity, such as H, have large energy-saving potentials (about up to 6% of heating load in our subjects). There are also some subjects whose sensitivities are also high although they have moderate energy consumption, such as D. They also have larger energy-saving potentials by a slight compromise in comfort. Of course, there are also some subjects who consume lower energy and have lower sensitivity, such as A, B, I, etc. They have less energy saving potentials. Other user's analysis can be done when the data is available. The operation managers could provide personalized guideline to the users based on the results to achieve energy-efficient building environment operation in personal level. Note that these observations in personal energy saving potentials are related to the results in Table 1 but much more comprehensive. The user's optimal temperature and humidity are closely related to the three terms $t_a, P_a, R+C$, which implies that the results rely on m_1, m_2, m_3 together. Also, our proposed model is data-driven. Many factors are concluded into the estimated

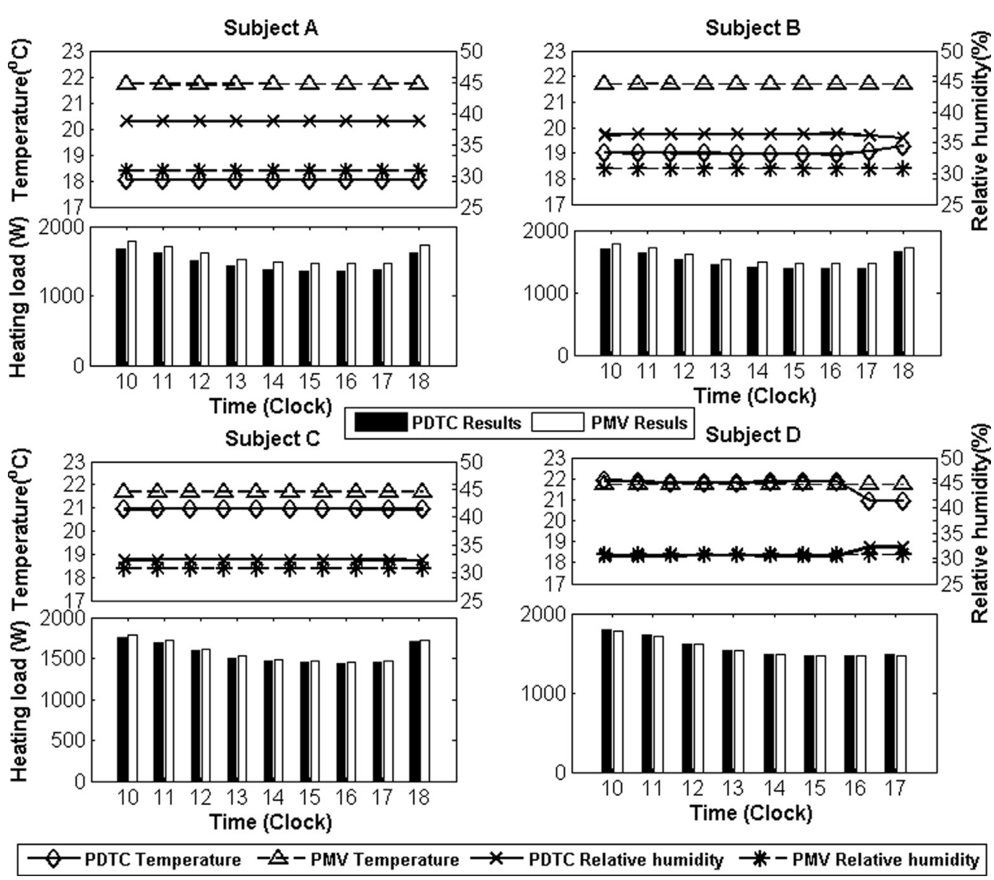

Fig. 11. Optimized set-points of temperature and humidity and corresponding heating loads for 4 typical subjects in one day. The upper parts of those figures show the set-point of indoor air temperature and relative humidity under PDTC constraint and PMV constraint. The lower part shows the heating load estimation during the period from the first voting moment of the day to the last voting.

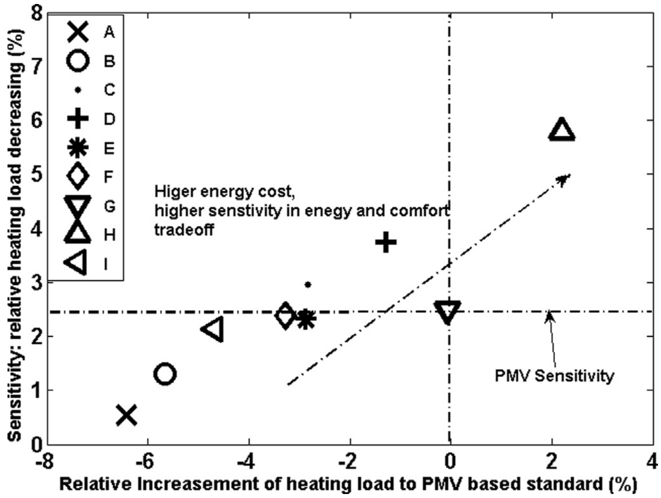

Fig. 12. Energy sensitivity analysis. The results were obtained when relaxing thermal comfort perception constrain form [−0.5, 0.5] to [−0.8, 0.8].

parameters and thus it is difficult to give specific physical explanation to the parameters. More sophisticated model is the future work.

5. Conclusion

In this paper, a new data-driven method describing the personalized dynamic comfort is proposed. The model is based on the heat balance of the human body with the unknown individual coefficients to identify. An application example of the model, personal energy saving analysis, is given. We utilize the relative energy saving ratio and sensitivity to identify personalized energy saving potentials in thermal acceptable range. The model and method were applied to an experiment in a real office room and a case study of personal energy saving analysis based on the experiment results and obtained satisfying results. The methods are expected to be of significant value in identifying personal comfort and energy saving potentials, and also in providing the personal control policy and feedback information of energy savings to users. By knowing each person's thermal comfort requirement, we can also control indoor environment better in a single occupant setting or settings with controllable micro-environment. The system for this model does not have specific requirement. Most of current HVAC systems, which control the air temperature and relative humidity, could use the proposed model, with the traditional thermostat replaced by the voting interface. The comfort set-point determined by the model will be the control target of the traditional HVAC controller. For multi-occupant environment with no controllable micro-environment, how to use our data-driven model is an open problem and needs further study. Other future work may include investigating personal dynamic comfort model in smaller time scale, studying the human behavior under the energy cost feedback information as well as evaluating the real effects of feedback information on energy savings, etc.

Acknowledgment

This work was supported by the Tsinghua-UTC Research Institute for Integrated Building Energy, Safety and Control Systems, and the United Technologies Research Center. Qianchuan Zhao received additional support from NSFC Grant (nos. 61074034, 60736027, 61021063, 61174072). The authors would also like to thank Dr. Alan Finn from United Technologies Research Center for the valuable comments and feedback. We also thank the anonymous reviewers for their helpful comments and suggestions that help to improve the quality of the paper.

References

[1] Fanger PO. Thermal comfort: analysis and applications in environmental engineering. Danish Technical Press; 1970.
[2] Hoof JV. Forty years of Fanger's model of thermal comfort: comfort for all? Indoor Air 2008;18:182–201.
[3] Humphreys MA, Nicol JF. The validity of ISO-PMV for predicting comfort votes in every-day thermal environments. Energy Build 2002;34:667–84.
[4] Gagge AP, Fobelets A, and Berglund L. A standard predictive index of human response to the thermal environment. ASHRAE Transactions 92.CONF-8606125-(1986).
[5] Gagge AP, Nishi Y. An effective temperature scale based on a simple model of human physiological regulatory response. ASHRAE Trans I 1971;77:247–62.
[6] ASHRAE. ASHRAE Standard 55-2010. ASHRAE; 2010.
[7] Stolwijk JAJ. Mathematical model of thermoregulation. Ann N Y Acad Sci 1980;355:98–106.
[8] Zhang H, Arens E, Huizenga C, Han T. Thermal sensation and comfort models for non-uniform and transient environments, part III: whole-body sensation and comfort. Build Environ 2010;45:399–410.
[9] de Dear RJ, Brager GS. Towards an adaptive model of thermal comfort and preference. ASHRAE Trans 1998;104:145–67.
[10] Humphreys MA, Nicol JF, Raja IA. Field studies of thermal comfort and the progress of the adaptive model. Adv Build Energy Res 2007;1:55–88.
[11] EN ISO 7730. Moderate thermal environments—analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort; 2005.
[12] YasuoKa A, Kubo H, Tsuzuki K, Isoda N. Interindividual differences in thermal comfort and the responses to skin cooling in young women. J Therm Biol 2012;37:65–71.
[13] Karjalainen S. Gender differences in thermal comfort and use of thermostats in everyday thermal environments. Build Environ 2007;42:1594–603.
[14] Haykin SS. Adaptive filter theory. 4th ed. Prentice Hall; 2002.
[15] U.S. Department of Energy. Building energy data book; 2012.
[16] Freire RZ, Oliveira GHC, Mendes N. Predictive controllers for thermal comfort optimization and energy savings. Energy Build 2008;40:1353–65.
[17] Atthajariyakul S, Leephakpreeda T. Real-time determination of optimal indoor-air condition for the thermal comfort, air quality and efficient energy usage. Energy Build 2004;36:720–33.
[18] Zhong K, Fu HM, Kang YM, Peng XF. Indoor thermal conditions and the potential of energy conservation of naturally ventilated rooms in summer, China. Energy Build 2012;55:183–8.
[19] http://dest.tsinghua.edu.cn/.
[20] http://apps1.eere.energy.gov/buildings/energyplus/.
[21] http://www.doe2.com/eQUEST/.
[22] Liu H, Zhao QC, Huang NJ, Zhao X. A simulation-based tool for energy efficient building design for a class of manufacturing plants. Automat Sci Eng IEEE Trans 2013;10:117–23.
[23] Nicol JF, Raja IA, Allaudin A, Jamy GN. Climatic variations in comfortable temperature: the Pakistan Projects. Energy Build 1999;30:261–79.
[24] Ljung L. System identification theory for the user. 2nd ed. Prentice Hall PTR; 2002.
[25] Kreider JF, Rabi A. Heat and cooling of buildings, design for efficiency. McGraw-Hill, Inc.; 1994.
[26] Haines RW, Wilson CL. HVAC systems design handbook [M]. McGraw-Hill; 2003.
[27] Lu YQ, etc. Air-conditioning design manual. 2nd ed. China Building Industry Press; 1995.
[28] Chu C, Jong T. Enthalpy estimation for thermal comfort and energy saving in air conditioning system. Energy Convers Manag 2008;49:1620–8.